Human Capital Externalities

Broersma, Lourens; Edzes, Arjen J. E.; Van Dijk, Jouke

Published in:
Regional Studies

DOI:
10.1080/00343404.2015.1053446

Publication date:
2016

Human Capital Externalities:

Effects for Low Educated Workers and Low Skilled Jobs

Abstract

Investments in human capital are essential themes in many policy programs. Besides the direct private returns of education, there is evidence of positive human capital externalities at the level of regions and firms. Our results show that both production and consumption externalities have positive effects on wages. Production externalities are transmitted at the level of firms and not at the regional level. For workers in low skilled jobs, consumption externalities dominate production externalities. Workers on low skilled jobs earn higher wages when working in co-operation with workers in high skilled jobs, while for low educated workers such co-operation with high educated workers is negative.

JEL: J21, J24, J31, R23

1. Introduction

Many western countries, regions and cities increasingly deal with the issue of how to improve the economic position of their low skilled and low educated citizens. The OECD (2006: 24) argues there are three reasons behind this attention. First, many countries and regions experience skill gaps and shortages in specific industries: '…there is a strong pressure to upgrade the skills of low-qualified workers on the domestic market so that they can fill vacancies for more qualified jobs and fuel economic growth'. Second, the desire to increase productivity for many countries, because '…higher productivity improves the position of firms on the global market, attracts inward investment and sustains job creation'. Productivity growth also helps to sustain the current level of welfare, which is important with the rise in ageing in many countries (Broersma *et al.*, 2014). Third, there is concern about the 'working poor', a '…vast category of workers in low-paid employment involving harsh working conditions'. This is related to the goal to realize an inclusive labour market in which outsiders (low productive people), who are now unemployed or inactive, can participate and get access to jobs. This is related to the shares of low educated and low skilled employment in the economy.

We address three themes that play a central role in today's discussions about improving the economic position of workers with a low education or of workers on low skilled jobs. The first theme concerns the underlying mechanisms that Human Capital Externalities (HCEs) have on employment and wages. In economics, externalities are issues that affect other people who are not part of a specific economic exchange. In our case, these externalities are related to human capital and how different levels in human capital may influence each other in their joint output level. Economists have long recognized that interactions among workers may help augment human capital that determines employment, wages and hence individual productivity. Several previous studies, including Rauch (1993), Acemoglu and Angrist (2000), Moretti (2004a) and Liu (2007, 2013), have tested this hypothesis by estimating the effect of city or state average educational attainment on individual earnings. Combes *et al.* (2012) study the wage and skill distribution by employment density in over 300 French employment areas. They find that both more high *and* more low skilled workers concentrate in dense areas rather than in less dense areas.

Heuerman *et al.* (2010) conclude in their literature review that studies of HCEs are still far from reaching a consensus, or, as Psacharapolous *et al.* (2002: 4) mentioned, '…the results overall are inconclusive'. HCEs are the increases in total earnings resulting from a one-year increase in education (Acemoglu and Angrist, 2001). Besides the direct private returns for the individual, there is evidence of positive human capital externalities or positive social rates of return in regions and firms (Schlitte, 2012). However, the exact nature of these externalities is unclear. Is it related to only a number of specific occupations of high or low skilled workers, as Wheeler (2001), Adamson et al. (2004) and Lee (2010) show, or is it related to the broader concept of creativity, as argued by Florida (2012, 2002)? Is it related to human capital of *workers* as argued by Lucas (1988), or to human capital of *residents*, as argued by Sassen (2001)? We will refer to this first type of social return, via workers, as production externality and the second type of returns, via residents, as consumption externality of human capital. In other words, production externality is defined here as the social rate of return on education of *workers* in a particular area. Workers in such an area benefit from a rise in education of an individual

worker, since all can learn from the higher education that this worker has acquired and they can all hence improve their output. To distinguish it from this production externality in an adequate way, we define consumption (as opposed to production) externality here as the social rate of return on education of *inhabitants* in a particular area[1]. It implies that when the level of education (or skills) of the inhabitants rises, their income is likely to increase and they are themselves thus more likely to hire persons with a lower education or a lower skill level to perform simple tasks on their behalf, like cleaning, maintenance and small repairs, gardening, shopping, baby-sitting and the like. In fact, this definition is close to what Acemoglu (1996) defines as pecuniary externality. This paper provides empirical evidence about the relative importance of these production and consumption externalities of human capital.

The second theme points to today's discussion on whether it is the education of the employee or the skill level of the job that determines economic progress. Education and skill are two different sides of the same medal. Education is a characteristic of a person and hence part of labour supply, while required skills are a characteristic of a job and hence part of labour demand[2]. Education and skill are not the same, as is simply illustrated in an increasing research literature on education or skill mismatch (Hamersma *et al.*, 2013; OECD, 2012, 2011; Desjardins and Rubenson, 2011; Quintini, 2011). Our purpose here is to assess whether the effects of HCEs differ when we look at either education or skill level.

The third and final theme points to discussions about the level at which HCEs influence wages of low educated employees in general or specifically workers on low skilled jobs. Is it the regional level or the firm level that matters most as mechanism for transferring HCEs? Moreover, as far as the distribution of different workers or jobs is concerned, do low educated workers or workers on low skilled jobs benefit, in terms of earning higher wages, from vicinity of many or of only a few high educated workers or workers on high skilled jobs? This distribution of education or skills within a region or firm is also a subject of our analysis. Basically it gives an indication of the possibility of co-operation between workers of different levels of education or the possibility of co-operation between (workers on) jobs of different levels of skill.

We proceed as follows. Section 2 presents a brief literature review of human capital externalities. Section 3 is about the methodology and data for The Netherlands that we use in our study. Section 4 presents, discusses and interprets the estimation and test results and finally section 5 concludes.

2. Human capital externalities: what mechanisms are at work?

Without a doubt human capital investments through formal education are an essential theme in many national, regional and local public policy programs. It is widely accepted that, at the individual level,

[1] Note that this definition of consumption externality is not the same as externality through consumption amenities (like restaurants, theaters and so on). Nor does it refer to gains in individual utility that goes with large or skilled cities. The latter is sometimes referred to technological externality (Gleaser *et al.* (2001)). Instead, our definition merely serves as a clear (and kind of opposite) distinction from the more usual production externality.

[2] We ignore that skill can also be seen as a characteristic of a person when it concerns his or her competences to fill a job.

more education leads to a better economic position in terms of higher wages and lower probability of becoming unemployed (Heuerman *et al.*, 2010, Moretti, 2004a, 2004b, 2012). The private internal rate of return to education is around 5-15%, i.e., an additional year of education of a worker leads to a 5-15% increase in his or her wage rate. The existence of high rates of private returns to education provides an incentive for individuals to invest in education. However, the benefits of education may not be solely attributed to the individual, but may spill-over to others as well. So the gains to an individual investing in education (the private return) could also raise the returns to the economy as a whole (the social return). This justifies public support for education. The literature on social rates of return to education focuses on the often cited fact that high educated countries, regions or cities generate more employment and higher labour productivity than low educated areas, resulting in higher wages for *all* inhabitants of that country, region or city (Berry and Glaeser, 2005; Moretti, 2004a, 2004b, 2012; Glaeser *et al.*, 1995).

HCEs reflect production externalities, in the sense that high educated workers increase the wage rate of other, lower educated, workers. More formally, production externalities raise these wages through a supply effect: extra employment for low educated workers is created because high educated workers share knowledge with low educated ones. Hence, when low educated workers interact with better-educated ones, they become more productive and creative and can hence earn higher wages. There is a longstanding research line that is engaged in investigating these HCEs (see Lucas, 1988, and overviews in Heuerman *et al.* 2010, Psacharopolous *et al.*, 2002; Moretti, 2010). At the same time, HCEs may also reflect, what we call, consumption externalities; i.e., the spending power of high educated or high skilled inhabitants (not necessarily only workers) has a positive effect on the income of inhabitants with a lower education or on a lower skilled job (Sassen, 2001). Sassen argues that this 'consumption externality' of educated individuals is oriented towards services often provided by individuals with a lower education or on a lower skilled job. Consumption externalities raise the income of these low educated through a demand effect: extra employment requiring low skills is created by demand from high educated individuals. The (wage) income of those low educated individuals will therefore rise. As a result, production externalities typically occur at the employment location, whereas consumption externalities occur at the residential location.

In addition, we have to determine whether there exists overlap between our definitions of production and consumption externalities. It is of course possible that a high educated worker, from whom a fellow (low educated) worker learns at the working location, will also hire (low educated) persons at his living location to perform certain tasks in and around the house, such as cleaning, gardening, shopping, baby-sitting and the like. Of course when someone lives in the same area where he or she works, that person can be present in both types of externalities. The question then is whether this is a problem of double counting or not. Both these types of externalities are distinguished separately and both have separate definitions, so the overlap of one person being in both types of externalities should not pose a problem as their definitions do not depend on each other. Moreover, even if this would be a problem, it still depends on the size of the regions we use. The smaller the surface, the larger commuting will be and for commuters there is no problem at all, since they have separate working and living areas to begin with. Of course now the actual size of the region is important. At the municipality

level, commuting is higher than at the (larger) regional NUTS-3 level. Still, commuting in The Netherlands is substantial even at NUTS-3, where about one third of the workers live in another region. At the municipality level even more than half of the workers commutes in from other municipalities where they live. Our regional demarcation in section 3 is based on 2-digit zip-code areas in the Netherlands, which amounts to 90 areas, that are somewhere in between municipalities (443) and NUTS-3 regions (40), so commuting will still be substantial and double counting in production and consumption externalities will not be a problem.

Besides these two forms of externalities, we also have to define what we mean by human capital. The levels of education and skill are intrinsically the two major components of human capital. Education is a characteristic of a person and is related to the qualifications and knowledge acquired through formal education. Skill, on the other hand, is a requirement of a job and is related to competences and expertise, acquired through experience and training a person needs to possess to be able to fill that job. As is the case with investment in physical capital, investment in human capital will only be undertaken by a wealth maximizing individual or a profit maximizing firm, if the expected return to this investment (the so-called net internal rate of return) exceeds the market rate of interest (Blundell *et al.* 1999).

Another interesting question is whether human capital spill-overs transmit between workers and firms. This relates to a large literature on the type of externalities that stimulate agglomeration, innovation and regional growth (for an overview: De Groot *et al.*, 2009) whereby a distinction is made in Porter-externalities, i.e. proximity of workers or firms within the *same* industry, and Jacobs-externalities, i.e. spill-over effects, as a result of proximity of firms in *different* industries. We have yet another approach. We focus on the spill-over effects on wage rates of individual workers from the distribution of workers with different educational levels in the region, in the work location (firm) and in the residential location.

Three important questions will be addressed in this paper. First, do production and consumption types of externalities exert an effect on the individual wage rate of workers and if so, which of these two externalities has the strongest effect? And how do these externalities affect especially the labour market position of the vulnerable group of low educated workers or workers on low skilled jobs? In that sense, do production externalities have an effect on the productivity, and thus the wage rate and hence income of low educated workers or of workers on low skilled jobs? Or do consumption externalities have an effect on the creation of low skilled jobs and hence the wage rate and income of low educated workers or workers who fulfil these low skilled jobs? Second, we will investigate the role of HCE at the firm level compared to the regional level on the wage rate of low educated workers and of workers on low skilled jobs. The final question is whether the distribution of low and high educated workers or of workers on low and high skilled jobs within a firm leads to HCEs?

3. Methodology and data

We start from the simple Mincerian framework to assess the impact of educational attainment and a host of individual, firm-level and regional background variables on the individual's wage rate. The focus of this paper is on the effect of contextual factors, such as the education level of co-workers in the same firm and co-residents in the residential region of the worker, *controlling for* the education level of the individual worker, along with other individual, firm level and regional characteristics. This implies the following model specification

$$\log(w_{i,j,k,t}) = \boldsymbol{\alpha} + \boldsymbol{X_{i,j,k,t}\beta} + \boldsymbol{Y_{j,k,t}\gamma} + \boldsymbol{Z_{k,t}\delta} + \varepsilon_{i,j,k,t}, \quad (1)$$
$$i = 1,\ldots, N, j = 1,\ldots,J, k=1,\ldots K$$

where $w_{i,f,r,t}$ is the vector of the hourly wage rate for the i-th individual, working in the j-th firm that is located in the k-th region at time t. In fact we follow the description of the multilevel model (MLM) in Aslam and Corardo (2012). The intercept α represents the intercept term including fixed effects of industry, firm size and time. Inclusion of these fixed effects implies a model specification as if the variables of (1) are all centred with respect to these three items. The explanatory variables in matrix $\boldsymbol{X}$ all relate to each individual employee i that works in firm j, which is situated in region k at time t. Variables that enter $\boldsymbol{X}$ are the individual education level, individual work experience, gender and so on. Of each individual employee i, we also know the firm j he or she works in and we know in which region k each firm j is located. Furthermore, we have information on variables $\boldsymbol{Y}$ that differ at the firm-level, but are similar for each individual employee working within that particular firm. One such variable is the firm level education, which is the average education level of all employees working within that firm. Finally, we also have information about variables $\boldsymbol{Z}$ that differ at the regional level, but are similar for each employee in each firm that is located within that particular region. Such variables are for example the regional worker education level, the regional inhabitant education level and regional unemployment.

Note that region or firm-level fixed effects are not included, because in a repeated cross section, like the one we use, having a large amount of dummies (like worker, but even firm or regional level fixed effect dummies) means we end up with degrees of freedom problems. This is because for each year the data are based on different workers in possibly different firms in a different location. Another problem related to adding large amounts of dummies to the model is the fact that these dummies may also capture the effects of actual variables in $\boldsymbol{Y}$ and $\boldsymbol{Z}$ at firm and regional level, respectively, thereby making these variables redundant at the expense of a dummy variable, where the interpretation is much weaker. Finally, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ represent vectors of constant parameters.

Equation (1) distinguishes the effects of human capital to the individual wage rate at the individual, the firm and the regional level. In specification (1) we consider a multilevel regression of the wage rate for each i-th individual worker, working in the j-th firm that is located in the k-th region. Even though multilevel data structures also arise in longitudinal studies where an individual's responses over time

are correlated with each other, our data set is not longitudinal. Instead, it can be regarded as a large cross section with three different levels of aggregation, the individual, firm and regional level, hence a multilevel model. In order to be able to estimate a model with such a complex data structure, we apply the appropriate multilevel model (MLM) estimation method[3]. Multilevel models recognise the existence of such data hierarchies by allowing for residual components at each level in the hierarchy. In such cases multilevel models yield correct inferences, which standard OLS does not, and they are thus still able to highlight the interest of specific groups. See also Snijders and Berkhof (2008).

The data set we use to estimate model (1) is the Working Conditions Survey (WCS) from the Dutch Ministry of Social Affairs. This is an annual employer-employee matched data set covering 1995-2006. The WCS is a stratified survey among firms, in which a sample of employees working within each firm is also questioned. Each annual wave of the WCS comprises on average 37000 employees in about 2000 firms. The WCS is not a panel in the sense that firms and workers can be followed through time. As was mentioned before, it comprises a repeated set of cross sections of firms and associated workers for each year, *i.e.* a repeated cross section. What this in fact means is that over the entire time period covered, our data set can be seen as one large cross-section of 368 thousand employees in 16 thousand firms. One of the variables determines the location of the firm, and thus the work location of the associated employees in terms of a 2-digit zip-code. These 2-digit zip-codes divide The Netherlands in 90 areas, around the larger cities. This means that the number of observations N, J and K in equation (1) are N=368,500, J=16,000 and K=90[4]. Since these zip-codes refer to the location of the firm, it helps us to identify possible production externalities taking place at the work location. Table 1 gives some descriptive statistics of our data set. The Appendix provides additional information on the data.

Table 1. Descriptive statistics of the WCS

	Mean	St. dev.	Min.	Max
Individual characteristics				
Gross hourly wage (euro's)	12.36	5.88	1.59	87.47
Education (years)	13,74	2,86	8	19
Experience (years)	19,58	11,18	0	52
Female	34%			
Part-time (less than 36 hours a week)	13%			
Firm characteristics				
Average education in firm (years)	13,36	1,95	8	19
Regional characteristics				
Population density (population per km^2)	1127	1194	104	5770
Regional unemployment (share of unemployed in population 15-64)	0.054	0.021	0.015	0.144
Human Capital Externalities (in years)				
Average education of the firm work location	13,65	0,76	10,43	16,57

[3] We use Stata version 8 of 2013, which allows for estimation of multilevel models with maximum likelihood estimation. See Hox (1998) for more on multilevel modelling.
[4] We do mention that the actual number of observation on individuals N with which the model is estimated depends on the type of variables that enter the model, but generally amount to 368500.

Average education of regional work location, excl. firm of the worker	13,67	0,78	9,59	16,57
Average education of regional residential location aged 15-64.	14.65	0.42	13.59	15.85

The mainstream approach is to evaluate effects of human capital externalities to productivity growth in a production function set up, instead of considering its effect on wages. The wage rate a firm is able to pay does have a direct link to its productivity. The higher the productivity of a firm, the higher the wage rate this firm can afford to pay its employees. In that sense productivity and wages are connected. It is always very difficult to get a hold of adequate data that give both firm level productivity and worker level of variables, such as education or skill. The WCS data set we use does contain the wage rate paid to employees by the firm, but not the productivity of the firm itself from which this wage is paid. Nevertheless, these two are closely connected as we have seen. The sequel just considers the effects of various levels of explanatory variable on the individual wage rate.

Educated and skilled labour

In our analysis, we can relate years of formal education to the individual worker's wage rate in equation (1)[5]. The indicators of human capital externalities (HCEs) are also based on the years of formal education. Education is therefore our key variable. We distinguish effects of education in (1), at the individual, at the firm and at the regional level. As mentioned before, these different levels imply a multilevel estimation procedure should be applied. Our model does distinguish different aggregation levels at which certain variables will be observed. First, for each individual employee, his years of education can be obtained from the WCS.[6] Second, for each firm, the years of education of its employees can be obtained by averaging the years of education of its workers. Third, the years of education in the region (2-digit zip-code) in which the firm is located can be obtained in a similar way, i.e. by averaging the number of years of education of each firm located in the region from the WCS. Education at the work location implies a focus on HCEs in terms of production externalities. In order to distinguish between production and consumption externalities, we have calculated the average years of education of the population between 15 and 64 at their residential location in each of the 90 2-digit zip-code areas of the Netherlands. These data are from Statistics Netherlands' Labour Force Survey (LFS), which we have added to the WCS.[7] This LFS is surveyed at the residential address and hence it measures the average education level at the place of living. This enables us to identify possible consumption externalities at the residential location, in contrast to production externalities that take place at the work location. Note that the number of years of education at the work location from the

[5] Educational attainment is classified in low, medium and high education. Low educated are defined as persons whose highest educational attainment is primary (ISCED 0, 1) and secondary education, first stage (ISCED 2). Intermediately educated have at attained at least higher and post-secondary education (ISCED 3, 4). High educated attained at least education of a tertiary and above level. ISCED refers to the International Standard Classification of Education.

[6] The definition and calculation of variables is described in the Appendix.

[7] Average education level is available for all 443 municipalities existing in 2007 for the total period 1996-2007 from various issues of Statistics Netherlands' Labour Force Survey (LFS). This municipal information is next rearranged to the same regional zip-code level as the one already used in the WCS.

WCS is slightly lower than that of the LFS at the residential location. This is due to the fact that certain industries are underrepresented in the WCS, like public administration and education, which causes a downward bias in the level of education in the WCS compared to the LFS. The LFS is a representative sample. It contains information on the education levels of the total population between 15 and 64 living in each of the 2-digit zip-codes in the Netherlands.

In the literature on skills it is common to equate a workers' skill to her or his education (Bacolod *et al*, 2009, 2010). However, education and skill are not the same. Although education is our key variable, we do take account of occupational skill-levels when interpreting the effect of HCEs on the rates of return on wages. The reason for this is that the group of low educated is not necessarily the same as the group occupying a low skilled job, because some low educated workers acquire additional competences allowing them to find a job at a higher skill level than their formal education allows. In fact, they are under-qualified for the job they fulfil. We distinguish between elementary, low, intermediately, high and scientifically skilled occupations.[8]

Distribution of workers by education and skill

The more high educated employees work together with low educated employees, the more likely it is that information spills over from these high to low educated workers. The same holds for the transfer of skills between workers on jobs of different skill levels. This means that low educated workers or workers on low skilled jobs may benefit when working next to high educated workers or workers on high skilled jobs. The WCS provides information per individual worker on the education level of that worker, as well as the skill level of the job that this worker fulfils. So, for each firm and each region we can identify the distribution of high and low educated workers and of workers on high and low skilled jobs.

This distribution is defined as $d = (e_{low} - e_{high})$, where $e_k = \sum_{i \in f} e_{i,k}/e$ is the share of workers/jobs in firm f by education/skill level k, where k refers to either low or high educated workers or to low or high skilled jobs. The rationale behind this distribution measure is to assess whether firms specialise in either low educated or high educated workers, or in terms of jobs, specialise in either low skilled or in high skilled jobs, to achieve productivity growth and hence being able to pay higher wages. This can be exemplified by two well-known companies: is it a MacDonald's-type of company, with many low skilled jobs (low educated workers) and just a few high educated or high skilled ones, or is it a Microsoft-type of company, with many high skilled jobs (high educated workers) and a few low educated or low skilled ones, where wages are affected positively. We specifically look at the wages of low educated workers or workers on low skilled jobs. The closer d is to zero, the more evenly workers of different skills or education are spread within a firm. A negative value of d means there are more high educated workers or workers on high skilled jobs in the firm than low educated workers or workers on low skilled jobs, i.e. a Microsoft type firm. A positive value of d, on the other hand, means

[8] The skill types in the WCS stem from a specific classification used by the Ministry of Social Affairs, which we have converted into the Standard Occupational Classification (SOC) of Statistics Netherlands. This SOC distinguishes occupations in elementary, low skilled, intermediately skilled, and high skilled and scientifically skilled.

more low skilled jobs (low educated workers) than high skilled (high educated) ones, so a MacDonald's type firm.

4. Results

As a first step, we estimate model (1) to assess the impact of production and consumption externalities of human capital to the wage rate of all workers, irrespective of their levels of education or skill. Next, we split-up the production externalities in externalities at the firm level and at the regional level net of the firm. This shows whether production externalities are present at the firm or at the regional level.

Table 2 presents the estimation results when equation (1) was applied to the WCS data base for the period 1995-2006 to which information was added about the average education level of the residents of age 15 through 64, at their place of residence. All model specifications of table 2 include 10 year dummies, implying that (1) is basically estimated as one large cross-section, as the dummies eliminate effects of time. Furthermore, all specifications include 15 industry dummies and 8 firm size dummies to account for industry and firm size effects. The typical variables commonly included in these types of Mincerian models, like experience and experience squared, show the expected sign and significance. The same holds for our finding that female workers have a lower hourly wage rates than average (and thus males). The fact that wages are higher in more densely populated areas and lower in areas with high unemployment is also in line with expectations and with results found in many other studies. Given our hierarchical multilevel method, the standard goodness of fit indicator, R^2, cannot be calculated, so we have added the value of the LR-test relative to a standard linear model. These LR-test values of the models of table 2 can mutually be compared in the sense that the lower the value of the LR-test, the better the model fits.

Table 2 comprises five different specifications to identify the effects of HCEs of persons working or living in a region, i.e., a 2-digit zip-code area, on the individual wage rate. Table 2 shows estimated coefficients and associated *z*-statistics, which have the same function as the usual *t*-statistics but now follow a standard normal distribution. Columns 2 and 3 of table 2 show the model with solely production externalities. We assume that besides the private returns to education for the worker, there is only regional production externality. The results imply that an extra year of education has a private rate of return of 7.8% on the individual wage. In fact, next to variables reflecting worker properties (like experience, gender and working time), this individual education level has the strongest effect on the individual wage rate in terms of the *z*-value. On top of this individual education effect, there is also a positive HCE of 0.3% from the average education of workers within the same region, i.e., 2-digit zip-code. In column 3, this zip-code education effect is split up into two parts. Firstly, a firm-level HCE effect, i.e., the level of education that other workers in the same firm have on the individual wage rate. Secondly, a regional HCE effect of the level of education of other workers in the same region (2-digit zip-code), excluding the firm. While column 2 of table 2 does show significant regional production externalities to the individual wage, in column 3 the regional production externalities are no longer significantly different from zero, but instead the now included firm-level externalities show up to be highly significant. This indicates that a worker does learn from other

workers inside the same firm, but not from other those working outside of the firm, but in the same region. An extra year of education of workers in the firm will raise the individual wage rate by 0.9% and thus the firm level production externality is larger than the regional level production externality of column 2 in table 2. Hence, knowledge transfer takes place on the work floor rather than in the region.

Column 4 of table 2 shows the results of a model which includes only consumption externalities. An extra year in average education of the inhabitants in a region will raise the individual wage rate in that region by 1.6%. This can be interpreted as an extra demand effect of higher educated inhabitants living in the region consuming more personal low educated or low skilled services that leads to the creation of additional jobs. In column 5 of table 2 both production and consumption externalities enter the model and we find the effect of regional production externalities is significantly smaller than that of regional consumption externalities. Do note that the LR-test for this model is smaller than the one of column 2. This is because of the strong effect that firm level externalities have on the individual wage rate. The final column of table 2 presents the estimation results of the model with production externalities, being again split up in a firm level effect and a regional effect, and consumption externalities. The regional production externality net of the firm is, just like column 2, insignificantly different from zero, but the firm-level production externality still has a similar coefficient and still is strongly significant. The regional consumption is slightly lower but is also still significant. These findings of a fairly constant values of firm-level production and regional consumption externality in all model specifications of table 2 implies that these HCE effects are quite robust. The size of the effect of production externality is slightly lower than that of the consumption externality. This final model of table 2 is preferred in terms of the lowest LR-test for goodness of fit compared to the LR tests for the other models.

We do note that certain coefficient values of table 2 are extremely small. Therefore we also estimated a model based on the same variables net of their mean and a model with standardized variables, *i.e.* net of mean and divided by the standard deviation.[9] Obviously the coefficient values change, but the associated *z*-values are very similar to the ones of table 2. Therefore we only report the estimations of the unadjusted variables of the model specifications in table 2, even though some may be small.Finally, the variables under the heading of ´variance components´ show how large the share of the total variance of the model is at which level of aggregation. We can clearly conclude that most of the error variance is at the individual level, than the firm level and finally the least at the regional level.

[9] These estimation results are available upon request.

Table 2.Estimation results of equation (1) with maximum likelihood, multilevel mixed effects model estimation, 1995-2006

Dependent variable	Log of hourly wage rate for all employees									
Model	coefficient	(*z*-value)	coefficient	(*z*-value)	coefficient	(*z*-value)	coefficient	(*z*-value)	coefficient	(*z*-value)
Individual worker level										
education level of individual	0.078	(482.8)	0.077	(451.7)	0.078	(484.1)	0.078	(482.8)	0.077	(451.7)
experience	0.044	(341.5)	0.044	(341.7)	0.044	(341.5)	0.044	(341.5)	0.044	(341.7)
experience squared	-7.1E-04	(-242.6)	-7.1E-04	(242.9)	-7.0E-04	(242.6)	-7.0E-04	(-242.6)	-7.1E-04	(-242.8)
female	-0.068	(-69.3)	-0.068	(-69.9)	-0.068	(-69.3)	-0.068	(-69.3)	-0.068	(-69.9)
part-time	0.195	(99.1)	0.193	(98.5)	0.195	(99.2)	0.195	(99.1)	0.193	(98.5)
Firm level										
average education of workers in firm			0.009	(18.3)					0.009	(18.3)
Regional level										
average education of workers in region	0.003	(2.14)					0.003	(2.01)		
average education of workers in region excl. firm			-8.7E-04	(-0.62)					-0.001	(-0.74)
average education of inhabitants aged 15-64 in region					0.016	(2.54)	0.015	(2.39)	0.014	(2.24)
regional population density	2.1E-05	(6.42)	2.1E-05	(6.48)	1.9E-05	(6.00)	1.8E-05	(5.68)	1.9E-05	(5.75)
regional unemployment rate	-0.512	(-4.39)	-0.523	(-4.48)	-0.521	(-4.55)	-0.516	(-4.50)	-0.526	(-4.57)
Controls										
time dummies (10)	Yes		Yes		Yes		Yes		Yes	
industry dummies (15)	Yes		Yes		Yes		Yes		Yes	
firm size dummies (8)	Yes		Yes		Yes		Yes		Yes	
number of variables	38		39		38		39		40	
number of observations	368,541		368,439		368,541		368,541		368,439	
Goodness of fit										
LR test vs. linear regression	65490		64514		65038		65032		64057	
Variance components										
2-digit zip-codes	5.7E-04		5.8E-04		4.9E-04		4.8E-04		5.0E-04	
firm-level	0.015		0.015		0.015		0.015		0.014	
all residuals	0.047		0.047		0.047		0.047		0.047	

The LR, or likelihood ratio, test rejects the null hypothesis if the value of this statistic is too small. The LR tests for the various models can be compared and the smaller the LR the better the fit. In that sense the model of the final column will be our preferred model specification. The variance components refer to each of the three aggregation levels in the multilevel model.

Externalities for employees with low education and employees on elementary and low skilled jobs

Table 3 shows the results for employees with a low education in columns 2 and 3 and the results for employees on an elementary or low skilled job are in columns 4 and 5. Columns 2 and 4 start with the same specification as the last column of table 2 and we will first discuss these results. Next, we add the distribution of low and high educated workers and of workers on low and high skilled jobs.

The private returns to education for workers with a low education and for workers on an elementary or low skilled job in table 3 is positive and about the same, with a rate of return of about 3.5%. So an additional year of education for workers with a low education or on a low skilled job will raise their wage rate with about 0.035. Note this is less than half the return we found for all employees regardless of education or skill in table 2. The coefficients for experience, experience squared, part-time work and population density are largely the same for low educated and low skilled and are also similar to what we found in table 2 for all employees. Next, we will discuss the differences between the models with all employees and of the low educated ones or the ones with an elementary or slow skilled job. Starting with the HCEs, we find interesting differences.

The consumption externality of low educated workers living in the region is about the same as that for all employees as reported in table 2, but is only weakly different from zero, while this effect is much larger - and significant - for workers on elementary and low skilled jobs. In line with the findings of table 2, table 3 also shows no significant effect of production externality at the regional level net of the firm. So it is only at the firm level where an exchange of ideas and help in production between workers takes place and not in the region outside the firm. Furthermore, such production externalities at the firm level turn out to be stronger for low educated workers than for workers on elementary and low skilled jobs. Compared to table 2 for all employees, we can also see that the coefficient values for production externality are about twice as large for low educated workers. For workers on elementary and low skilled jobs they are about 50% larger. So they profit more from a rise in education or skill at the firm level than all employees do.

Table 3 also shows that inclusion of the distribution variable, $d = (e_{low} - e_{high})$, has significant but opposite effects in the model of low educated workers compared to the one of workers on elementary and low skilled jobs. For low educated workers we find a positive coefficient for the distribution d of low minus high educated employees, whereas it is negative for the distribution of elementary and low minus high and scientific jobs. So as far as education of employees is concerned, firms specialising in high educated workers have $d<0$, so the wage rate for its low educated workers – table 3, column 3 – will fall. When instead firms have more low educated workers than high educated ones, $d>0$ so the wages for low educated workers will rise. Notice however that at the same time table 3 shows a remarkable change in the effect of production externalities. This implies that workers inside the firm benefit stronger from rise in education when d is added. So for firms with many high educated workers, this compensates for the lower effect that d has on the wage rate, and likewise but than the other way around for firms with many low educated workers.

Table 3. Estimation results of equation (1), of employees with a low education and employees on elementary and low skilled jobs, 1995-2006 (t-statistics between brackets based on clustered standard errors)

Dependent variable	Log of hourly wage rate for							
	low educated employees				employees on elementary and low skilled jobs			
Model	Coefficient	(*t*-value)	coefficient	(*t*-value)	coefficient	(*t*-value)	coefficient	(*t*-value)
Individual worker level								
education level of individual	0.035	(82.4)	0.035	(81.2)	0.034	(96.3)	0.035	(97.9)
experience	0.046	(297.2)	0.046	(297.2)	0.045	(264.2)	0.045	(264.5)
experience squared	-7.4E-04	(-225.7)	-7.4E-04	(-225.7)	-7.5E-04	(-198.9)	-7.5E-04	(199.2)
female	-0.051	(-40.6)	-0.050	(-39.6)	-0.014	(-10.1)	-0.013	(-10.0)
part-time	0.206	(92.6)	0.205	(92.1)	0.176	(76.8)	0.174	(76.0)
Firm level								
average education of workers in firm	0.019	(28.7)	0.025	(27.9)	0.013	(17.7)	0.002	(1.62)
Regional level								
average education of workers in region excl. firm	-0.001	(-0.79)	-0.001	(-0.86)	-0.003	(-1.61)	-0.003	(-1.67)
average education of inhabitants aged 15-64 in region	0.012	(1.87)	0.012	(1.98)	0.021	(3.48)	0.019	(3.03)
regional population density	1.4E-05	(4.65)	1.5E-05	(4.93)	1.7E-05	(6.16)	1.6E-05	(5.97)
regional unemployment rate	-0.377	(-3.17)	-0.392	(-3.26)	-0.509	(-4.10)	-0.470	(-3.77)
Distribution at firm-level of								
low vs. high educated workers			0.040	(10.7)				
elementary plus low vs. high plus scientifically skilled jobs							-0.073	(-17.4)
Controls								
time dummies (10)	Yes		Yes		Yes		Yes	
industry dummies (15)	Yes		Yes		Yes		Yes	
firm size dummies (8)	Yes		Yes		Yes		Yes	
number of variables	40		41		40		41	
number of observations	188,532		188,532		131,773		131,773	
Goodness of fit								
LR test vs. linear regression	33357		33238		246992		24172	
Variance components								
2-digit zip-codes	3.5E-04		3.8E-04		2.4E-04		2.5E-04	
firm-level	0.013		0.013		0.015		0.015	
all residuals	0.034		0.034		0.033		0.033	

The LR tests for the various models can be compared and the smaller the LR the better the fit. The variance components refer to each of the three aggregation levels in the multilevel model.

We should however be aware of the fact that the education level says nothing about the job requirements, which in fact determine the wages being paid by the firm. The current economic crisis lead many high educated workers to accept jobs of a lower skill level and so earn the – lower – wages that go with that type of job. It might therefore actually be better to consider the skill level of the job that a worker has, rather than his level of education. The last two columns of table 3 show the model for the wage rate of workers on jobs of an elementary and low skill level. The model in the fourth column of table 3 shows the similar model specification as in the last column of table 2, but now no longer for all workers, but only for those on elementary and low skilled jobs. The final column of table 3 shows that inclusion of the distribution variable of elementary plus low skilled jobs minus high plus scientifically skilled ones, has a negative significant effect. This means that in firms that specialize in high skilled jobs (i.e. the Microsoft type of firm), the wage rate for their workers on elementary and low skilled jobs is higher. In other words, workers on low skilled jobs benefit from possible learning effects of workers on high skilled jobs and as a result they earn higher wages. Reversely, when a firm has more workers on elementary and low skilled jobs than on high and scientifically ones (i.e. the McDonald's type of firm), this will exert a negative effect on the wage rate of elementary and low skilled jobs. Also not that in this case there is a similar kind of compensatory effect in the firm-level education on the wage of elementary and low skilled jobs. Whereas the third column shows a positive significant production externality of 1.3%, when d is added in the final column, it fell to a mere 0.2% and is no longer significantly different from zero. Clearly this effect is now in the distribution-variable of the final model. For firms who specialize in high educated workers, this leads to an additional increase in wages of those with low skilled jobs of 7.3%.

5. Conclusions

In this paper we have investigated three themes that play a central role in today's discussions on how to improve the economic position of low educated workers or workers on low skilled jobs. First, we unravel the debate on human capital externalities (HCEs) and ask ourselves the question if consumption or production externalities occur and if so, which one of the two has the strongest effects. Our results show that an extra year of education has a very strong private rate of return of almost 8% on the individual wage rate. On top of that, employees in general benefit from HCEs, i.e. workers in a firm or region benefit from the average education level in their firm or region. These HCEs can be split-up in both production and consumption externalities of human capital and both have significant positive effects on the wages of individual employees, but the magnitude of the effect of regional consumption externalities is larger than the production externalities at the firm level.

Second, when we take a closer look at the production externalities by making a distinction between externalities taking place at the regional level or at the firm level, it becomes clear that production externalities are only transmitted at the level of firms and not at the regional level. Clearly, for learning effects to take place, there needs to be some kind of vicinity on the work floor.

Third, especially for workers on low skilled jobs it matters in which type of firm they work in order to benefit from externalities due to the presence of workers with high skilled jobs. It appears that the

distribution of workers on low and high skilled jobs within a firm has a negative wage-effect for low skilled workers when a firm has many workers on low skilled jobs and only a few on high skilled (MacDonald's type of firm). In contrast to that, the distribution has a positive wage effect when a firm has many high skilled jobs and only a few low skilled ones (Microsoft type of firm). However, when instead we add the education distribution within a firm to the model for the low educated workers, we found the exact opposite. In other words, now a McDonald's type of firm has a positive wage effect instead of the Microsoft type of firm. Do note however, that even though the education of workers and the skill level of jobs will largely coincide, many low educated workers try to get a higher skilled job when they can, but also high educated workers settle for less when times are bad. So there is a substantial difference between education and skill level.

We find that a policy strategy with the aim of attracting high educated workers is beneficial for enhancing the labour market position of low skilled and low educated workers. The effect mainly takes place via consumption effects. So, investing in amenities that improve the residential attractiveness is a promising strategy. Stimulating production spill-overs via learning effects at the regional level does not work, because this mechanism works only at the firm level and is especially effective when workers on low skilled jobs in firms work together with many high skilled workers. So, from a learning point of view, aiming at higher productivity and wages of low skilled workers, attracting Microsoft type of firms is preferred above attracting McDonald's type of firms.

References

Acemoglu, D. (1996). A Microfoundation for Social Increasing Returns in Human Capital Accumulation. *The Quarterly Journal of Economics*, 111: 3, 779-804.

Acemoglu, D. and J. Angrist (2001). How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws. B.S. Bernanke and K.Rogoff (ed), *NBER Macroeconomics Annual 2000*, Vol. 15, Boston: MIT Press.

Adamson, D.W., D.E. Clark, and M.D. Partridge (2004). Do Urban Agglomeration Effects and Household Amenities Have a Skill Bias?, *Journal of Regional Science,* 44: 201–223.

Aslam, A. and L. Corrado (2012). The Geography of Well-Being. *Journal of Economic Geography,* 12, 3: 627-649.

Bacolod, M., B.S. Blum and W.C. Strange (2010). Elements of skill: Traits, intelligences, education, and agglomeration. *Journal of Regional Science*, 50, 1: 245-280.

Bacolod, M., B.S. Blum and W.C. Strange (2009). Urban interactions: soft skills versus specialization. *Journal of Economic Geography*, 9: 227-262.

Berry, C. and E. Glaeser (2005). The divergence of Human Capital Levels across cities. *Papers in Regional Science*, 84, 3: 407-444.

Broersma, L. I. Noback and J. van Dijk (2014), Labour productivity growth as panacea for ageing? - How The Netherlands can keep up its level of welfare. Mimeo: University of Groningen.

Canton, E. (2009). Human capital externalities and proximity: evidence from repeated cross-sectional data. *De Economist*, 157, 1: 79-105.

Combes, P., G. Duranton, L. Gobillon and S. Roux (2012). Soring and local wage and skill distributions in France. Regional Science and Urban Economics, 913-930.

Combes, P., G. Duranton and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63: 723-742.

Blundell, R., L. Dearden, C. Meghir and B. Sianesi (1999). Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy. *Fiscal Studies*, 20, 1: 1-23.

De Groot, H.L.F., J. Poot and M.J. Smit (2009). Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: Capello, R. and P. Bijkamp (2009) Handbook of regional growth and development theories

Desjardins R. and K. Rubenson (2011). *An analysis of skill mismatch using direct measures of skills*. Education Working Papers No. 63, Paris: OECD.

Florida, R., C. Mellander, K. Stolarick and A. Ross (2012). Cities, skills and wages, Journal of Economic Geography, 12: 355-377.

Florida, R. (2002) *The Rise of the Creative Class*, New York: Basic Books.

Glaeser, E.L., J.A. Scheinkman and A. Shleifer (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, 36: 117-143.

Glaeser, E.L., J. Kolko and A. Saiz (2001). Consumer city. *Journal of Economic Geography*,1, 27-50.

Hamersma, M., A.J.E. Edzes and J. van Dijk (2013). Underqualification as an opportunity for low educated workers. *Environment and Planning C: Government and Policy*. In print.

Henderson, J.V. (2007). Understanding knowledge Spill-overs. *Regional Science and Urban Economics,* 37: 497–508.

Heuerman, D., B. Halfdanarson and J. Suedekum (2010*)*. Human Capital Externalities and the Urban Wage Premium: Two literatures and their Interrelations. *Urban Studies*, 47, 4: 749-767.

Howells, J. (2012). The geography of knowledge: never so close but never so far apart. *Journal of Economic Geography*, 12: 1003-1020.

Hox, J. (1998). Multilevel modelling: when and why. In: I. Balardjan, R. Mathar and M. Schader (eds), *Classification, Data Analysis and Data Highways*, New York: Springer and Verlag, 147-154.

Liu, Z. (2007). The external returns to education: Evidence from Chinese cities. *Journal of Urban Economics*, 61: 542–564.

Liu, Z. (2013). Human capital externalities in cities: evidence from Chinese manufacturing firms. *Journal of Economic Geography* forthcoming.

Lee, S. (2010). Ability Sorting and Consumer City. *Journal of Labor Economics*, 68, 1: 20-33.

Lucas, R.E. (1988). On the Mechanisms of Economic Development. *Journal of Monetary Economics*, 22, 1: 3-42.

Möller, J. and Haas, A. (2003). The agglomeration differential reconsidered: an investigation with German micro data 1984–1997, In: Broecker, J., D. Dohse and R. Soltwedel (Eds) *Innovation Clusters and Interregional Competition*, Berlin: Springer.

Moretti, E. (2012). *The New Geography of Jobs*. Houghton Mifflin Harcourt

Moretti, E. (2004a). Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. *American Economic Review*, 94, 3: 656-690.

Moretti, E. (2004b). Human Capital Externalities in cities. Henderson, J.V. and J.F. Thisse. *Handbook of Regional and Urban Economics*, Volume 4. Elsevier.

OECD (2006). *Skills upgrading: new policy perspectives.* Paris: OECD

OECD (2011). *Employment Outlook 2011*. Paris: OECD.

OECD (2012). *Better Skills, Better Jobs, Better Lives. A Strategic Approach to Skills Policies*, Paris: OECD.

Psacharropoulos, G. and H.A, Patrinos (2002). Returns to Investment in Education: a further update. *World Bank Policy Research Working Paper* 2881.

Quintini, G. (2011). Over-qualified or under-skilled, A review of existing literature, OECD Social, Employment and Migration Working Papers No. 121, Paris: OECD.

Sassen, S. (2001). *The Global City: New York, London Tokyo*, Princeton University Press.

Schlitte, F. (2012). Local human capital, segregation by skill, and skill-specific employment growth , *Papers in Regional Science*, 91, 1: 85-106.

Snijders, T.A.B. and J. Berkhof (2008).Diagnostic checks for multilevel models. In J. de Leeuw and E. Meijer (eds), *Handbook of Multilevel Analysis*, New York: Springer, 141-175.Wheeler, C.H. (2001). Search, Sorting, and Urban Agglomeration. *Journal of Labor Economics*, 19: 880–898.

Appendix – Data sources and definitions

Tabel A1: The regional dimension of the data base

2-digit zip-code	Municipality name	Average years of education region	Observations of individuals	Observations of firms
10	Amsterdam1	14,76	16346	1131
11	Amsterdam2, Volendam, Zwanenburg, Amstelveen	14,39	10823	596
12	Hilversum, Laren, Huizen	14,06	2669	221
13	Almere, Weesp	13,27	2346	181
14	Bussum, Uithoorn, Purmerend	13,26	3674	395
15	Zaandam, Wormerveer	13,22	3723	380
16	Enkhuizen, Hoorn (NH)	13,24	2365	234
17	Heerhugowaard, Schagen, Den Helder	13,47	2661	316
18	Alkmaar	13,91	3055	260
19	Castricum, Beverwijk, IJmuiden	13,54	3992	271
20	Haarlem	14,33	3378	278
21	Heemstede, Hoofddorp, Lisse	13,83	5261	365
22	Noordwijk (ZH), Katwijk (ZH), Wassenaar, Voorschoten, Leidschendam, Voorburg, Rijswijk (ZH)	14,16	5756	465
23	Leiden	14,24	4421	313
24	Alphen aan den Rijn	12,93	2994	280
25	Den Haag	14,76	7752	470
26	Delft, Naaldwijk	13,74	5619	469
27	Zoetermeer, Waddinxveen	14,17	3585	272
28	Gouda	13,61	2701	238
29	Capelle aan den IJssel, Krimpen aan den IJssel, Alblasserdam, Ridderkerk	13,35	7531	559
30	Rotterdam	13,68	15230	933
31	Schiedam, Vlaardingen, Maassluis, Hoogvliet	13,88	7925	470
32	Spijkenisse, Hellevoetsluis, Middelharnis, Oud Beijerland	13,33	4020	411
33	Dordrecht, Zwijndrecht, Papendrecht	13,38	6827	512
34	IJsselstein, Nieuwegein, Woerden	14,12	6732	532
35	Utrecht	14,79	9294	480
36	Maarssen	13,76	2122	166
37	Zeist, Bilthoven, Soest, Barneveld	14,07	5876	520
38	Amersfoort, Harderwijk, Nijkerk(GLD)	13,69	7924	565
39	Veenendaal, Doorn, Driebergen-Rijsenburg	13,74	5771	503
40	Tiel	12,51	1390	100
41	Culemborg, Leerdam	13,39	2502	152
42	Gorinchem	13,43	2403	216
43	Schouwen Duiveland, Walcheren	13,49	2414	192
44	Noord and Zuid Beveland	13,41	1724	153
45	Zeeuws Vlaanderen	13,35	2406	173
46	Bergen op Zoom	12,93	2149	248
47	Roosendaal	13,18	4436	382
48	Breda	13,31	6289	454
49	Oosterhout (NB)	13,03	2752	194
50	Tilburg1	13,11	6203	415
51	Tilburg2, Waalwijk	12,79	3847	296
52	Den Bosch, Vught, Boxtel	13,55	8673	565
53	Zaltbommel, Oss	13,46	3236	257
54	Uden, Veghel	13,16	5351	407

55	Veldhoven, Valkenswaard	13,52	3595	277
56	Eindhoven	14,24	7988	492
57	Helmond, Deurne	13,65	4434	359
58	Venray	13,60	2861	189
59	Venlo	13,49	6523	422
60	Weert, Roermond	13,39	6058	454
61	Sittard, Geleen	13,88	5439	376
62	Maastricht	14,01	4794	342
63	Heerlen1	13,45	1462	145
64	Heerlen2, Brunssum, Kerkrade	13,46	4304	318
65	Nijmegen	13,72	4383	311
66	Wijchen, Elst (GLD)	12,92	3173	303
67	Wageningen, Ede (GLD)	14,10	3726	253
68	Arnhem, Oosterbeek, Velp	14,34	5934	404
69	Zevenaar, Dieren	13,43	3002	257
70	Doetinchem	12,89	3126	287
71	Winterswijk, Lichtenvoorde	13,35	2346	223
72	Zutphen	13,36	3243	205
73	Apeldoorn	13,61	5545	438
74	Deventer, Nijverdal, Goor	13,72	6008	469
75	Enschede, Hengelo (OV), Oldenzaal	14,03	6741	622
76	Almelo, Vriezenveen	13,36	3033	242
77	Dedemsvaart, Coevorden	12,51	2379	236
78	Emmen	13,13	1787	169
79	Hoogeveen, Meppel	12,81	3442	277
80	Zwolle, Nunspeet	13,58	4633	293
81	Raalte, Epe	13,10	1816	157
82	Lelystad, Kampen	13,09	2816	242
83	Emmeloord, Steenwijk	13,23	2073	203
84	Gorredijk, Heerenveen	13,38	2089	232
85	Joure	13,59	870	94
86	Sneek	13,55	772	81
87	Bolsward	12,81	691	73
88	Franeker	13,21	711	88
89	Leeuwarden1	13,80	2929	150
90	Leeuwarden2	13,57	548	68
91	Dokkum	12,83	896	76
92	Drachten, Bergum	13,25	2916	212
93	Roden	13,05	1004	88
94	Assen	13,39	2423	219
95	Stadskanaal	12,75	549	59
96	Hoogezand, Veendam, Winschoten	13,40	2260	185
97	Groningen1	14,16	3431	286
98	Groningen2	13,34	462	33
99	Appingedam, Winsum (GN)	13,58	1178	108

Definition of variables

Years of Education

The WCS education variable (v1014) distinguishes seven education levels, but does not specify the number of years of formal education. To calculate educational attainment expressed in education years, the following rules are applied: 8 years of education for employees who completed primary education (v1014=20); 12 years of education for employees who completed secondary general education (v1014=30), lower vocational education (v1014=31), or intermediate general education (v1014=40); 16 years of education for employees who completed intermediate vocational education (v1014=41); 18 years of education for employees who completed upper vocational education (v1014=50) and 18 or 19 years of education for employees who completed university education (v1014=60).

The full primary school program is 8 years. The first year is not compulsory, but most children do participate in the first year. We take account of a major reform of the Dutch higher education system in 1982 ("twee fasen structuur"), in which the nominal duration of a university program was shortened from 5 to 4 years. For example, we assume that employees with a university degree and older than 42 in 2004 have attended a 5 year academic program, while employees with a university degree and not older than 42 in 2004 followed a 4 year academic program (cf. Webbink, 2007). For the calculation of the fraction of skilled workers we include persons with upper professional education ("hoger beroepsonderwijs") and persons with university education ("wetenschappelijk onderwijs").

For the years of education of the employed labour force living in each of these 2-digit zip-codes, the same assumption is used, be it that the LFS only provides information at three levels of educational attainment: low, intermediate and high. For these two levels we have assumed 11, 14 and 18 years of education respectively.

Experience

Labor market experience is calculated as follows: $experience = age - education - 4$

where *education* is the number of years of formal education, and it is assumed that children start education at the age of 4.

Wages

The hourly wage rate is determined from

$whour = v22a \times 12 / v63a \times 52 - v67a \times v63a / 5$

where v22a is the monthly wage, v63a is the number of hours worked per week, v67a is the number of holidays ("roostervrije dagen").

Part-time Work

An employee is defined as a part-time worker if v63a/v66a<0.5, where v63a stands for the number of hours worked per week and v66a for the normal number of hours worked per week.

Sector Structure

Sector 1 : Agriculture and fisheries
Sector 2 : Mineral extraction
Sector 3 : Industry
Sector 4 : Energy- and water control companies
Sector 5 : Construction
Sector 6 : Reparation of consumer articles and trade
Sector 7 : Hotels, restaurants, cafés
Sector 8 : Transport, storage and communication
Sector 9 : Financial institutions
Sector 10 : Rental and business services
Sector 11 : Public management, mandatory social insurances
Sector 12 : Education
Sector 13 : Health and social services
Sector 14 : Culture, recreation and other services

Cleaning the Data

Persons with an age below 15 and above 65, and persons with experience below zero are dropped from the sample. Also, we dropped observations where the hourly wage is below 75% of the legal minimum wage or above 12 times the legal minimum wage. Minimum wages for employees under 23 are calculated as a percentage of the minimum wage for persons aged 23 and older,